POTENCY
PROBIT ANALYSIS

by

Richard J. Daum [1/]

and

Clyde Givens [2/]

United States Department of Agriculture

Biometrical Services

Beltsville, Maryland

May 1963

1/ Biometrician, Biometrical Services, USDA, ARS. Present address- M & T Chemical Inc., Rahway, New Jersey.
2/ Programmer, Biometrical Services, USDA, ARS, Beltsville, Maryland

This program obtains the potency (ratio of equally effective doses) and its 95% confidence limits for upwards of five parallel dose-response lines through a regression of log-dose on probit response employing the maximum likelihood procedure described by D. J. Finney in Probit Analysis (Cambridge University Press, 1952).

Input consists of a parameter card for the experiment followed by up to five subsets of data. Each subset of data is preceded by a subparameter card and each subset of data may contain from two to twenty data cards (doses). The parameter card for the experiment consists of three 4-digit numbers which identify the "experiment", a dummy number identifying the first subset of data, and the number of subsets of data in the experiment. In addition to this, the first parameter card also contains two Chi square values (for testing for homogeneity of the data and parallelism of the regression lines respectively), the value of Student's "t" for setting confidence limits for heterogeneous data, the critical value of F_{05} for testing significance of the regression, and a "Log Factor" for making all doses greater than unity before transforming dose to $\log_{10}$(dose). Each subset of data that follows (up to five) must be preceded by a sub-parameter card.

Each sub-parameter card contains three 4-digit numbers which identify the experiment, the subset of data to which the sub-parameter card belongs, the number of data cards (doses) composing the subset, the number of animals that were observed, and the number that responded in each untreated control (check). The use of sub-parameter cards permits a different check for each subset of data.

Each data card contains two 4-digit numbers identifying the experiment and the subset to which the data belong (for ease of sorting and ordering of the data cards), followed by three 6-digit fields which contain the dose used, the number of animals treated, and the number that responded.

The program computes the percent response in the untreated control, with provision for zero (0) observations and /or zero response in the untreated control. The program, as it reads each data card, converts the dose to $\log_{10}$(dose) and the numbers observed and responded to a proportion, adjusting each proportion for the untreated control by use of Abbott's Formula. The adjusted (net) responses are then transformed to probits by use of a polynominal approximation. (Hastings 1955, Approximations for Digital Computers). The program then types out the numbers identifying the experiment, the subset to which the observation belongs, the dose administered, the net response (as a percent) and the proportion that responded in the untreated control. After typing out this information the weighted regression of log-dose on probit response is computed using an iterative procedure (maximum likelihood) with the restriction that the dose-response lines for all subsets are parallel.

Iteration continues until the last and next to last slope values (B's) agree within ±.00005 or until 20 cycles have been completed. After cycling until convergence the program types out the numbers identifying the experiment and subset, the number of doses in each subset, the weighted and corrected (for weighted mean) sums of squares of log-dose (X), probit-response (Y), cross products (XY), and the sum of squares attributable to regression for each subset (dose-response line). The program continues, again typing out the number identifying the experiment, the number of doses in each subset, the weighted means of log-dose and probit-response, the sum of the weighting coefficients for each subset and the intercept of each weighted probit regression line.

Following this the program next presents a summary of the experiment in an analysis of linear regression format, which includes the degrees of freedom (computed), sum of squares (total, regression, deviations from regression, and parallelism), and pertinent parameters (F ratio needed, F ratio obtained, Chi

square values used in the tests for homogeneity and parallelism). Parallelism is the difference between the sum of squares attributable to regression resulting from fitting individual regression lines to each subset of data and from fitting a single regression line (the same slope) to all subsets of data. This definition of parallelism, because it is an over simplification, is not entirely correct as stated, and the reader interested in the exact procedure is referred to either the accompanying flow sheet or to D. J. Finney's Probit Analysis.

Following the presentation of the analysis of linear regression, the program then test for homogeneity of the data before typing out the number of iterations completed (first line obtained from a weighted regression of log-dose on probit-response using the observed proportions that responded in the weighting coefficients and is not counted in determining the number of iterations), g (the precision of estimating the slope which enters in setting confidence limits about the potencies), the critical value of t_{05} used in computing the confidence limits about the potencies, the standard error, the standard error of the slope, and the slope of the regression lines (single slope for all regression lines). If the data are heterogeneous, t_{05} is set equal to Student's t_{05} with degrees of freedom associated with the deviations mean square (total degrees of freedom minus the number of lines, where the total degrees of freedom is the sum of the number of observations in the experiment minus one for each dose-response line within the experiment), and the square root of the deviations mean square is set equal to the standard error. If the data are homogeneous $t_{05} = 1.96$ and standard error = 1.00 are used in computing the 95% confidence limits of the potencies. The standard error of the slope (STDEB) is invariable $\sqrt{(\text{standard error})^2/SS_{xx}}$.

The program then test for parallelism. If the lines are not parallel LINES NOT PARALLEL is typed out, and the program proceeds without pause to read the parameter card for the next experiment, bypassing the remaining calculations

in the program. If the lines are parallel the program continues with the remaining calculations, testing next for significance of the regression.

If the regression is not significant, as tested by the critical value of F_{05}, NON SIG REGRESSION is typed out and the program proceeds without pause to read the parameter card for the next experiment. If the regression is significant the program continues, computing and typing out the LD-30, LD-50, LD-70 and LD-90 for each regression line (subset of data). The program then proceeds to compute and type out potencies (ratios of LD-50 of first line divided by LD-50 of each succeeding dose-response line) and their 95% confidence limits. The program then proceeds without pause to read the parameter card for the next experiment, which initiates another series of calculations for fitting parallel dose-response lines using the probit analysis procedure outlined by D. J. Finney in Probit Analysis.

The weighting coefficients and the procedures used in the program are similar to those used in a Probit Analysis program written by R. J. Daum, Clyde Givens and Gary Bearden (Biometrical Services, USDA, Beltsville, Md.) for fitting simple, individual, dose-response lines using the Maximum Likelihood procedure described by D. J. Finney (ibid). The reader seeking greater detail on these points is referred to this program. Running time with this program is approximately four minutes with typewriter output. If it is anticipated that more than an occasional set of data will be processed by this program it is recommended that a card-punch output be used. This change may be accomplished by simply changing the PRINT statements to PUNCH and recompiling the program.

The advantages of this program are (1) a neatly labeled and well organized print-out of sufficient information to permit plotting the original and computed dose-response lines, and to spot the subset(s) of data which resulted in either biological or statistical invalidity if such occurs, (2) the inclusion of numerous checks which permits the analysis to be completed only on data which are both

statistically and biologically valid, and (3) the use of procedures which permit an indefinite number of analyses to be performed without additional instructions or interference from the machine operator.

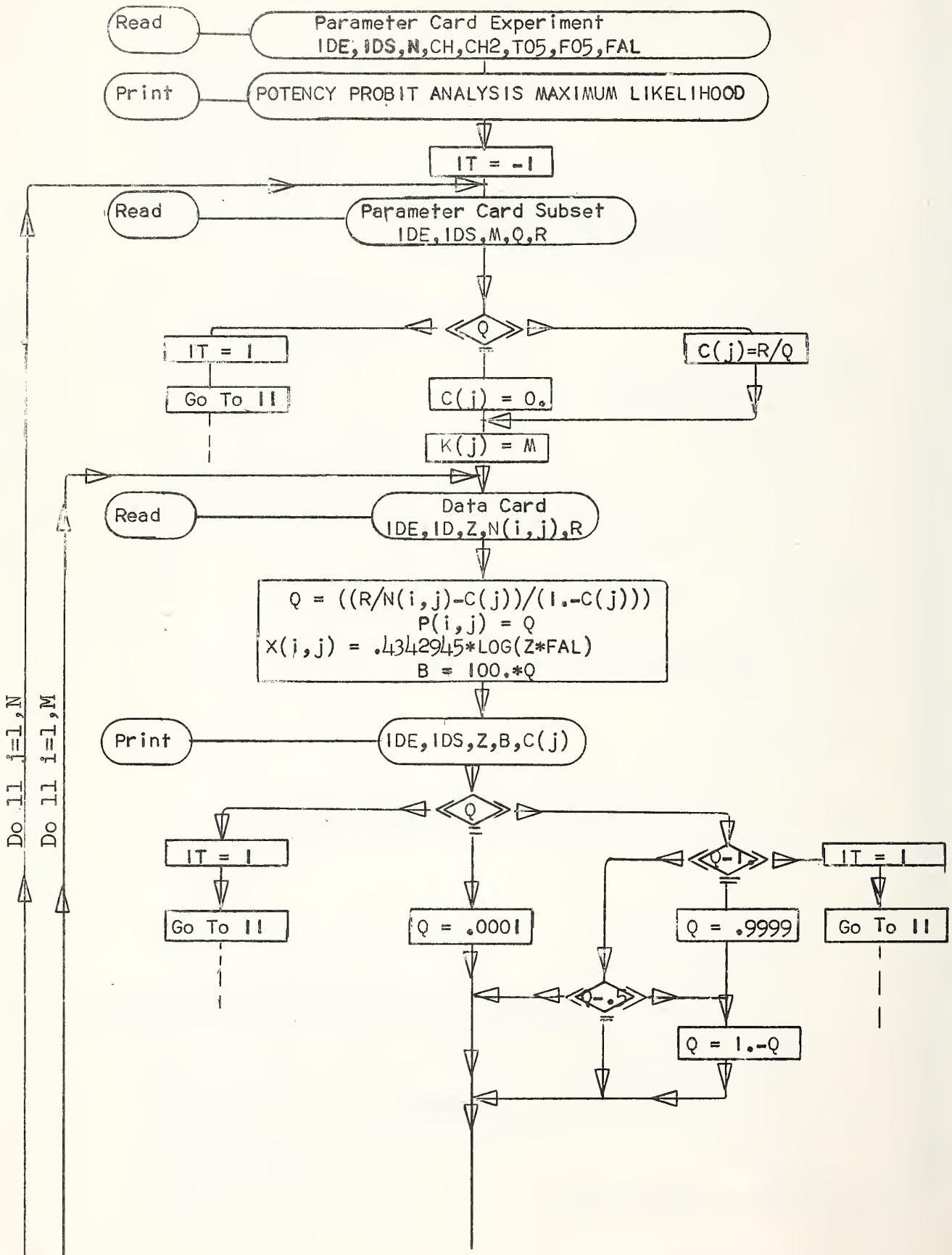
```
07300 C      POTENCY PROBIT ANALYSIS MAXIMUM LIKELIHOOD, 1620 40K, MULTIPLE   6
07300 C      FORMAT FORTRAN, AUTO DIVISION. THIS PROGRAM OBTAINS THE WEIGHTED
07300 C      LINEAR REGRESSION OF PROBIT RESPONSE ON LOG DOSE BY THE MAXIMUM
07300 C      LIKELIHOOD PROCEDURE DESCRIBED BY D.J. FINNEY IN PROBIT ANALYSIS
07300 C      (CAMBRIDGE UNIVERSITY PRESS).  POTENCY IS THE RATIO OF THE FIRST
07300 C      LD-50 DIVIDED BY THE LD--50 OF EACH SUCCEEDING SET OF DATA IN
07300 C      SAME ORDER AS DATA ARE READ INTO COMPUTER.  MAXIMUM OF 5 SETS
07300 C      OF 20 DOSES EACH.
07300 C      AUTHORS R.J. DAUM AND CLYDE GIVENS, BIOMETRICAL SERVICES, USDA,
07300 C      BELTSVILLE MARYLAND, MAY 1963.
07300        DIMENSION C(5),K(5),VN(20,5),X(20,5),P(20,5),WN(20,5),A(5),XB(5)
07300        DIMENSION YB(5),SX(5),SY(5),SXY(5),SWN(5),Y(20,5)
07300     49 FORMAT (2I4,3F6.0)
07342     50 FORMAT (43H POTENCY PROBIT ANALYSIS MAXIMUM LIKELIHOOD,///)
07468     52 FORMAT(22H IDENT  SET       DOSE,10X,12HNET RESPONSE,7X,5HCHECK)
07632     53 FORMAT (12HSET NO.DOSES,6X,3HSSX,13X,3HSSY,13X,4HSSXY,12X,5HSSREG)
07854     51 FORMAT (/)
07876    54 FORMAT(12HSET NO.DOSES,6X,4HXBAR,12X4HYBAR,12X,3HSNW,9X,9HINTERCEPT)
08098     55 FORMAT (/30H ANALYSIS OF LINEAR REGRESSION/)
08192     56 FORMAT (20H SOURCE VARIATION DF,11X,2HSS,14X,2HMS,)
08342     57 FORMAT (18H TOTAL             ,I4,F16.6,16X,F13.3,6H   F05)
08476     58 FORMAT  (18H REGRESSION       ,I4,2F16.6,F13.3,6H  FCAL)
08576     59 FORMAT (18H DEV REGRESSION    ,I4,2F16.6,F13.3,8H  CHI SQ)
08680     60 FORMAT (3I4,5F6.0)
08738     62 FORMAT (2I5,3F16.6)
08780     63 FORMAT (2I5,4F16.6)
08828     64 FORMAT (I3,5F12.6)
08876     70 FORMAT (18H PARALLELISM       ,I4,2F16.6,F13.3,8H  CHI SQ)
08980     71 FORMAT (3H IT,8X,1HG,9X,3HT05,8X,4HSTDE,8X,5HSTDEB,7X,5HSLOPE)
09196     72 FORMAT (18HNON SIG REGRESSION)
09256     73 FORMAT (18HLINES NOT PARALLEL)
09316 74FORMAT(10HIDENT  SET,9X,4HLD30,12X,4HLD50,12X,4HLD70,12X,4HLD90)
09538 75FORMAT(10HIDENT  SET,4X,11HUPPER LIMIT,7X,7HPOTENCY,7X,11HLOWER LIMIT)
09718        D1=7.0523078E-02
09742        D2=4.2282012E-02
09766        D3=9.2705272E-03
09790        D4=1.5201430E-04
09814        D5=2.7656720E-04
09838        D6=4.3063800E-05
09862      1 READ 60,IDE,IDS,N,CH,CH2,T5,F5,FAL
09970        PRINT 50
09994        PRINT 52
T0018        IT=-1
T0054        DO 11 J=1,N
T0066        READ 60,IDE,IDS,M,B2,B
T0138        PRINT 51
T0162         IF(B2)99,101,102
T0218    101 C(J)=0.
T0266         GO TO 103
T0274    102 C(J)=B/B2
T0334    103 G=1./(1.-C(J))
T0406        K(J)=M
T0454        DO 11 I=1,M
T0466        READ 49,IDE,IDS,Z,VN(I,J),Q
T0598        Q=((Q/VN(I,J))-C(J))*G
T0742        B=Q*100.
```

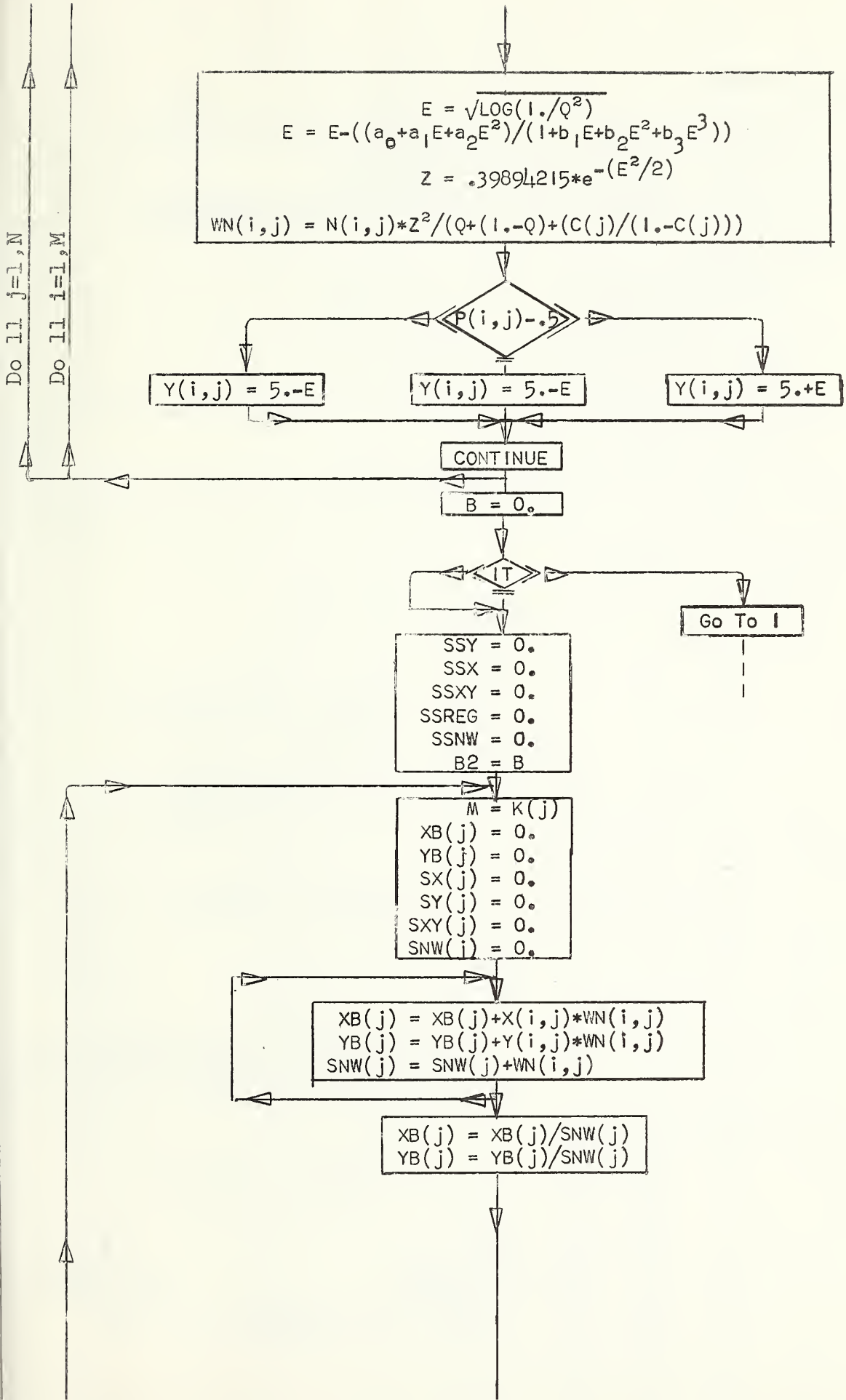
```
T0778          X(I,J)=.4342945*LOG(Z*FAL)
T0898        P(I,J)=Q
T0982        PRINT 62,IDE,IDS,Z,B,C(J)
T1078        IF (Q)99,3,4
T1134      3 Q=.0001
T1158        GO TO 8
T1166      4 IF (Q-1.)6,5,99
T1234      5 Q=.9999
T1258      6 IF (Q-.5)8,8,7
T1326      7 Q=1.-Q
T1362      8 E=SQRT(LOG(1./(Q*Q)))
T1434        B=2.515517+E*(E*.010328+.802853)
T1506        E=E-(B/(1.+(E*(E*(E*.001308+.189269)+1.432788))))
T1638        Z=.39894215*EXP(-E*E*.5)
T1722        WN(I,J)=(VN(I,J)*Z*Z)/(Q+((1.-Q)+C(J)*G))
T2022        IF(P(I,J)-.5)9,9,10
T2150      9 Y(I,J)=5.-E
T2246        GO TO 11
T2254     99 IT=1
T2278     10 Y(I,J)=5.+E
T2374     11 CONTINUE
T2446        PRINT 51
T2470        B=0.
T2494        IF (IT)12,1,1
T2550     12 SSY=0.
T2574        SSX=0.
T2598        SSXY=0.
T2622        SSREG=0.
T2646        SSNW=0.
T2670        B2=B
T2694        DO 15 J=1,N
T2706        M=K(J)
T2754        XB(J)=0.
T2802        YB(J)=0.
T2850        SX(J)=0.
T2898        SY(J)=0.
T2946        SXY(J)=0.
T2994        A1=0.
T3018        A2=0.
T3042        A3=0.
T3066        SWN(J)=0.
T3114        DO 13 I=1,M
T3126        B=WN(I,J)
T3210        A1=A1+B
T3246        A2=A2+B*X(I,J)
T3354     13 A3=A3+B*Y(I,J)
T3498        XB(J)=A2/A1
T3558        YB(J)=A3/A1
T3618        SWN(J)=A1
T3666        A1=0.
T3690        A2=0.
T3714        A3=0.
T3738        DO 14 I=1,M
T3750        B=WN(I,J)
T3834        Q=X(I,J)-XB(J)
T3954        E=Y(I,J)-YB(J)
```

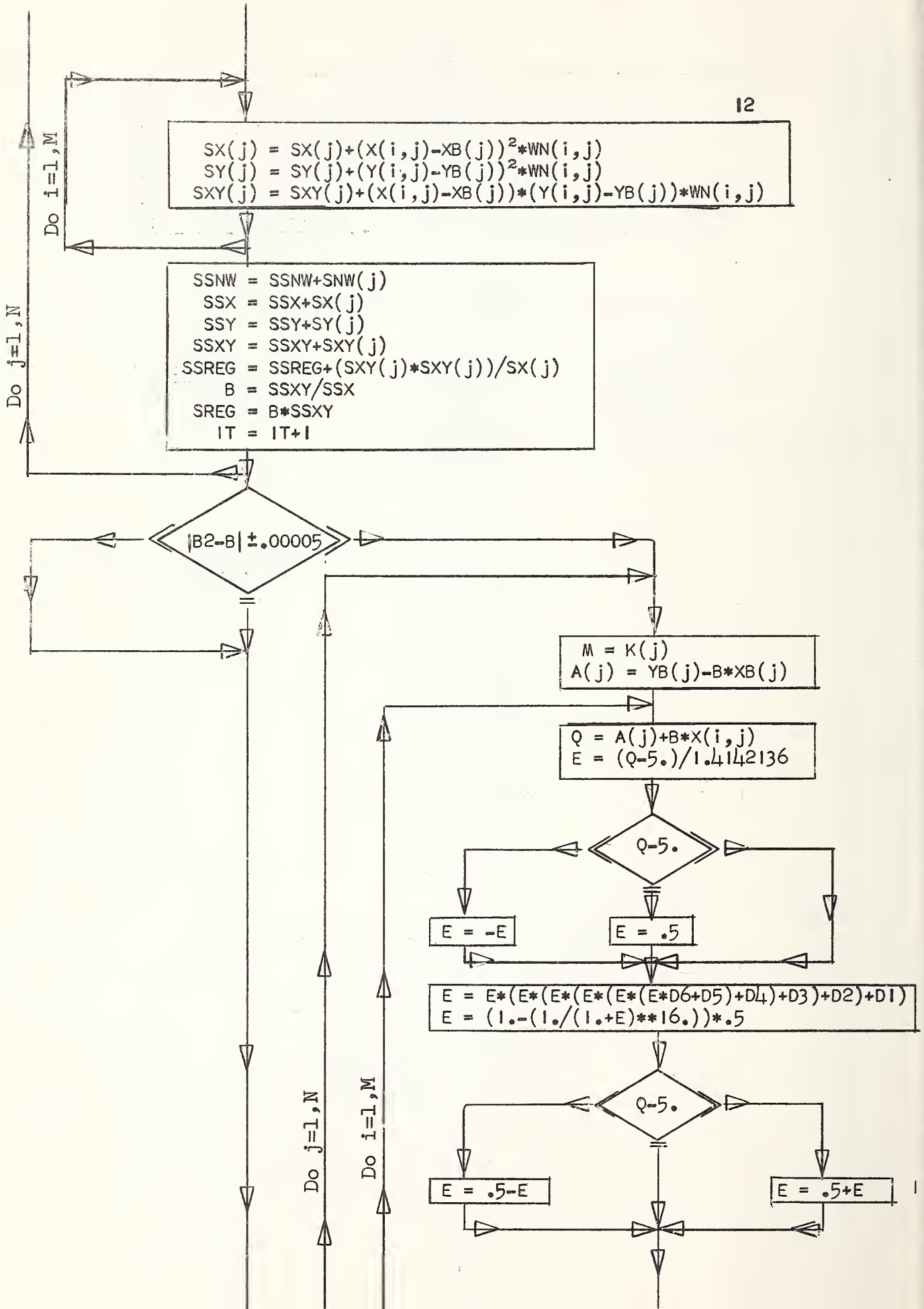
```
T4074          A1=A1+B*Q*Q
T4134          A2=A2+B*E*E
T4194       14 A3=A3+B*E*Q
T4290          SX(J)=A1
T4338          SY(J)=A2
T4386          SXY(J)=A3
T4434          SSNW=SSNW+SWN(J)
T4494          SSX=SSX+A1
T4530          SSY=SSY+A2
T4566          SSXY=SSXY+A3
T4602       15 SSREG=SSREG+A3*A3/A1
T4698          B=SSXY/SSX
T4734          SREG=B*SSXY
T4770          IT=IT+1
T4806           IF(IT-20)222,122,122
T4874      222 IF (B2-B)23,25,24
T4942       23 IF (B2-B+.00005)122,25,25
T5022       24 IF (B2-B-.00005)25,122,122
T5102      122 DO 22 J=1,N
T5114          G=1./(1.-C(J))
T5186          M=K(J)
T5234          A(J)=YB(J)-B*XB(J)
T5366          DO 22 I=1,M
T5378          Q=A(J)+B*X(I,J)
T5510          E=(Q-5.)/1.4142136
T5558          IF (Q-5.)16,17,18
T5626       16 E=-E
T5662          GO TO 18
T5670       17 E=.5
T5694          GO TO 21
T5702       18 E=E*(E*(E*(E*(E*(E*D6+D5)+D4)+D3)+D2)+D1)
T5858           E=(1.-1./(1.+E)**16.)*.5
T5954          IF (Q-5.)19,21,20
T6022       19 E=.5-E
T6058          GO TO 21
T6066       20 E=.5+E
T6102       21 Z=.39894215*EXP(-(Q-5.)*(Q-5.)*.5)
T6234          Y(I,J)=Q+((P(I,J)-E)/Z)
T6414       22 WN(I,J)=(VN(I,J)*Z*Z)/((1.-E)*(E+C(J)*G))
T6786          GO TO 12
T6794       25 PRINT 53
T6818          LTDF=0
T6842          SSPAR=SSREG-SREG
T6878          DO 26 J=1,N
T6890          A3=SXY(J)*SXY(J)/SX(J)
T7010          LTDF=LTDF+K(J)-1
T7082       26 PRINT 63,J,K(J),SX(J),SY(J),SXY(J),A3
T7298          LRDF=1
T7322          LPDF=N-1
T7358          LDDF=LTDF-N
T7394          A2=LDDF
T7430          A3=LPDF
T7466          C(4)=6.2816
T7490          C(3)=5.5244
T7514          C(2)=5.
T7538          C(1)=4.4756
```

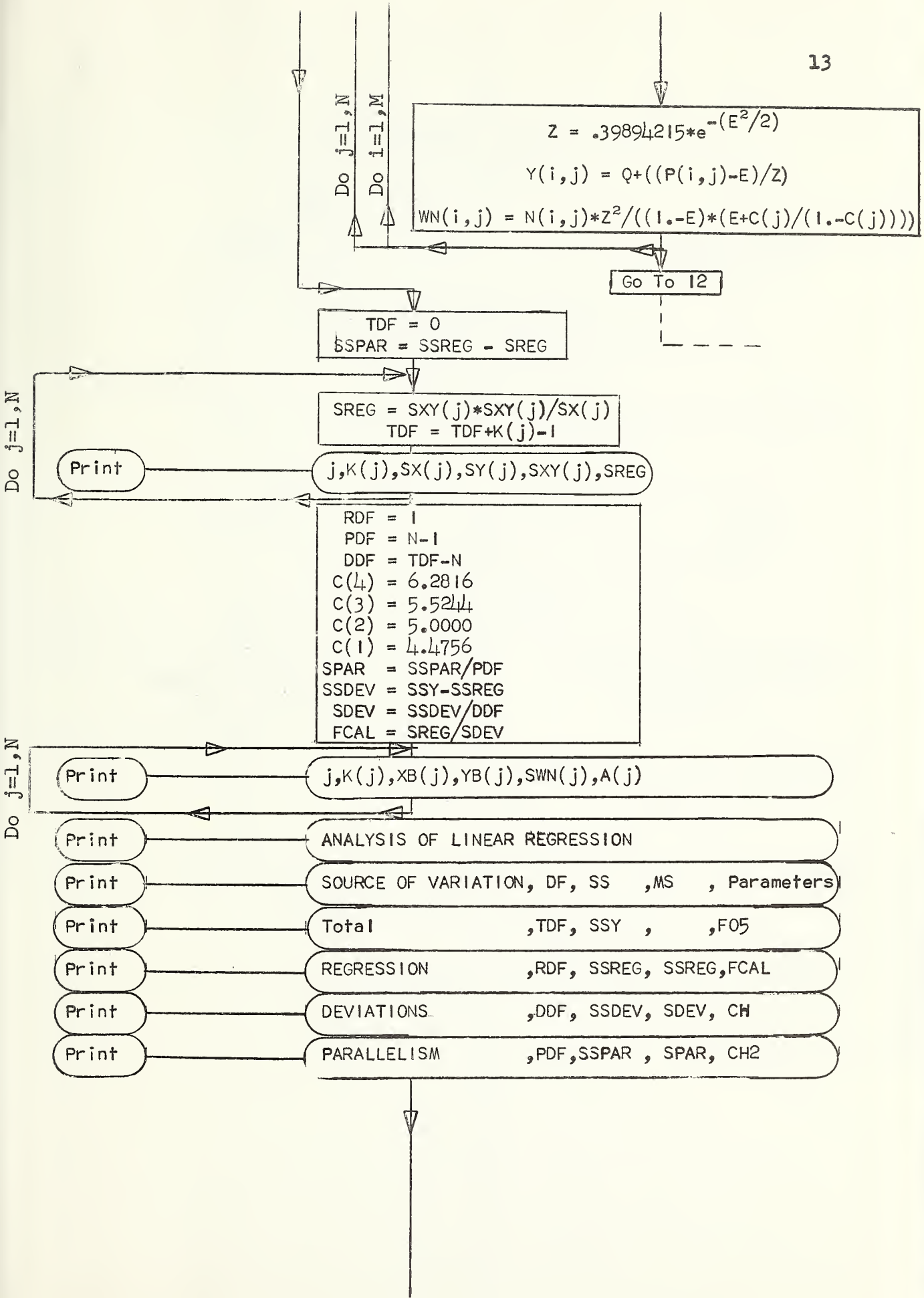
```
17562          SPAR=SSPAR/A3
17598          SSDEV=SSY-SSREG
17634          SDEV=SSDEV/A2
17670          FCAL=SREG/SDEV
17706          PRINT 51
17730          PRINT 54
17754          DO 27 J=1,N
17766       27 PRINT 63,J,K(J),XB(J),YB(J),SWN(J),A(J)
18006          PRINT 55
18030          PRINT 56
18054          PRINT 57,LTDF,SSY,F5
18102          PRINT 58,LRDF,SREG,SREG,FCAL
18162          PRINT 59,LDDF,SSDEV,SDEV,CH
18222          PRINT 70,LPDF,SSPAR,SPAR,CH2
18282          IF(SSDEV-CH)28,29,29
18350       28 Z=1.
18374          T5=1.96
18398          GO TO 30
18406       29 Z=SDEV
18430       30 SSXY=SQRT(Z/SSX)
18478          G=T5*T5*Z/SREG
18538          Z=SQRT(Z)
18562          PRINT 51
18586          PRINT 71
18610          PRINT 64,IT,G,T5,Z,SSXY,B
18694          PRINT 51
18718          IF (G-.05)31,32,32
18786       31 G=0.
18810       32 IF(F5-FCAL)34,34,33
18878       33 PRINT 72
18902          GO TO 1
18910       34 IF (CH2-SSPAR)35,36,36
18978       35 PRINT 73
19002          GO TO 1
19010       36 PRINT 74
19034          DO 38 J=1,N
19046          DO 37 I=1,4
19058       37 YB(I)=(10.**((C(I)-A(J))/B))/FAL
19262       38 PRINT 63,IDE,J,YB(1),YB(2),YB(3),YB(4)
19382          PRINT 51
19406          PRINT 75
19430          DO 39 J=2,N
19442          A1=(A(J)-A(1))/B
19514          Q=A1-XB(1)+XB(J)
19586          E=A1+(G*Q/(1.-G))
19682           A3=1.-G
19718          A3=((T5*Z)/(B*A3))*SQRT((A3/SWN(1))+(A3/SWN(J))+(Q*Q/SSX))
19994          POTUL=10.**(E+A3)
20066          POT=10.**A1
20102          POTLL=10.**(E-A3)
20174       39 PRINT 63,IDE,J,POTUL,POT,POTLL
20282          PRINT 51
20306          GO TO 1
20314          END
SW 1 OFF TO IGNORE SUBROUTINES
```

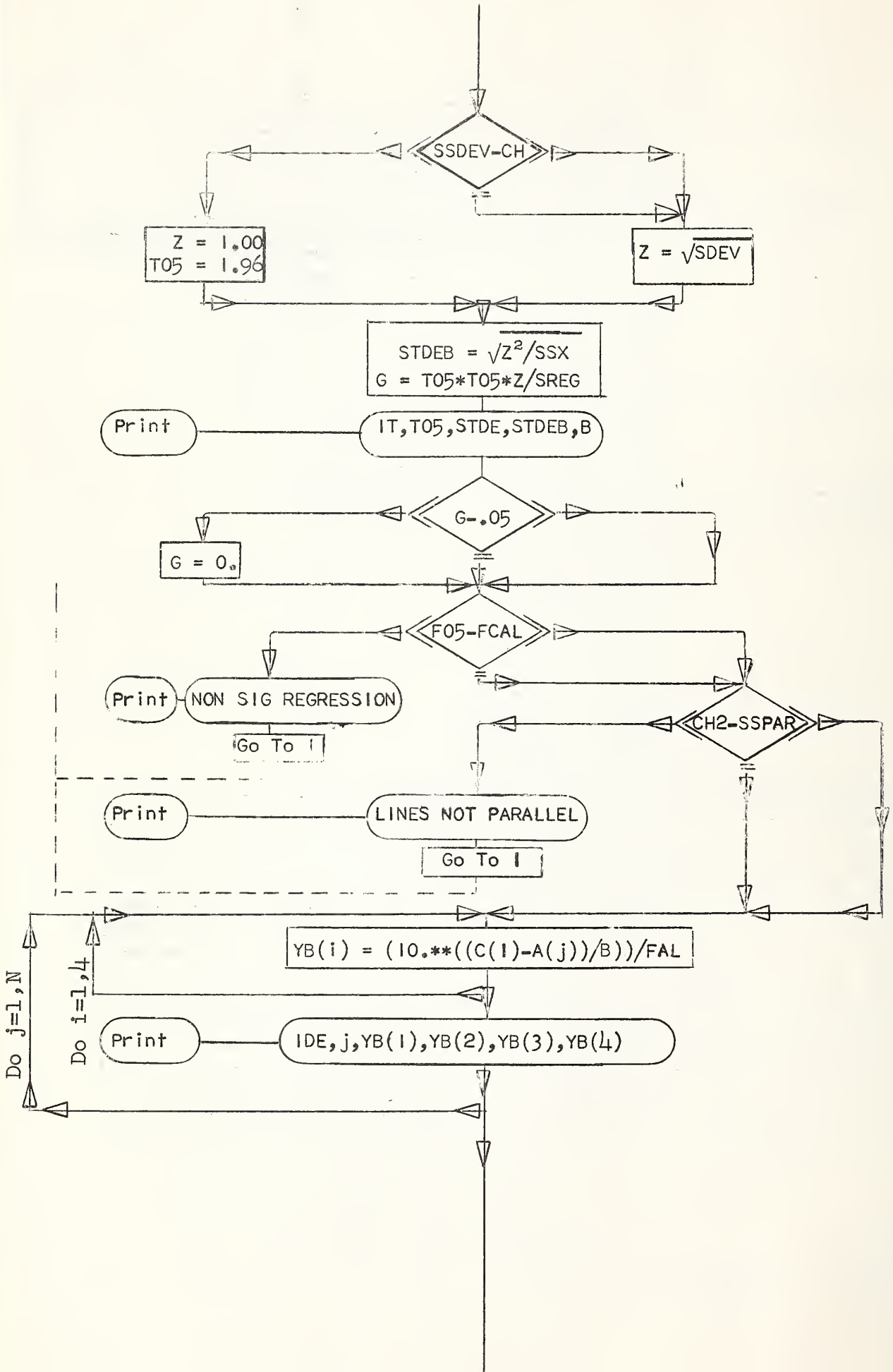
FLOW SHEET
POTENCY PROBIT ANALYSIS

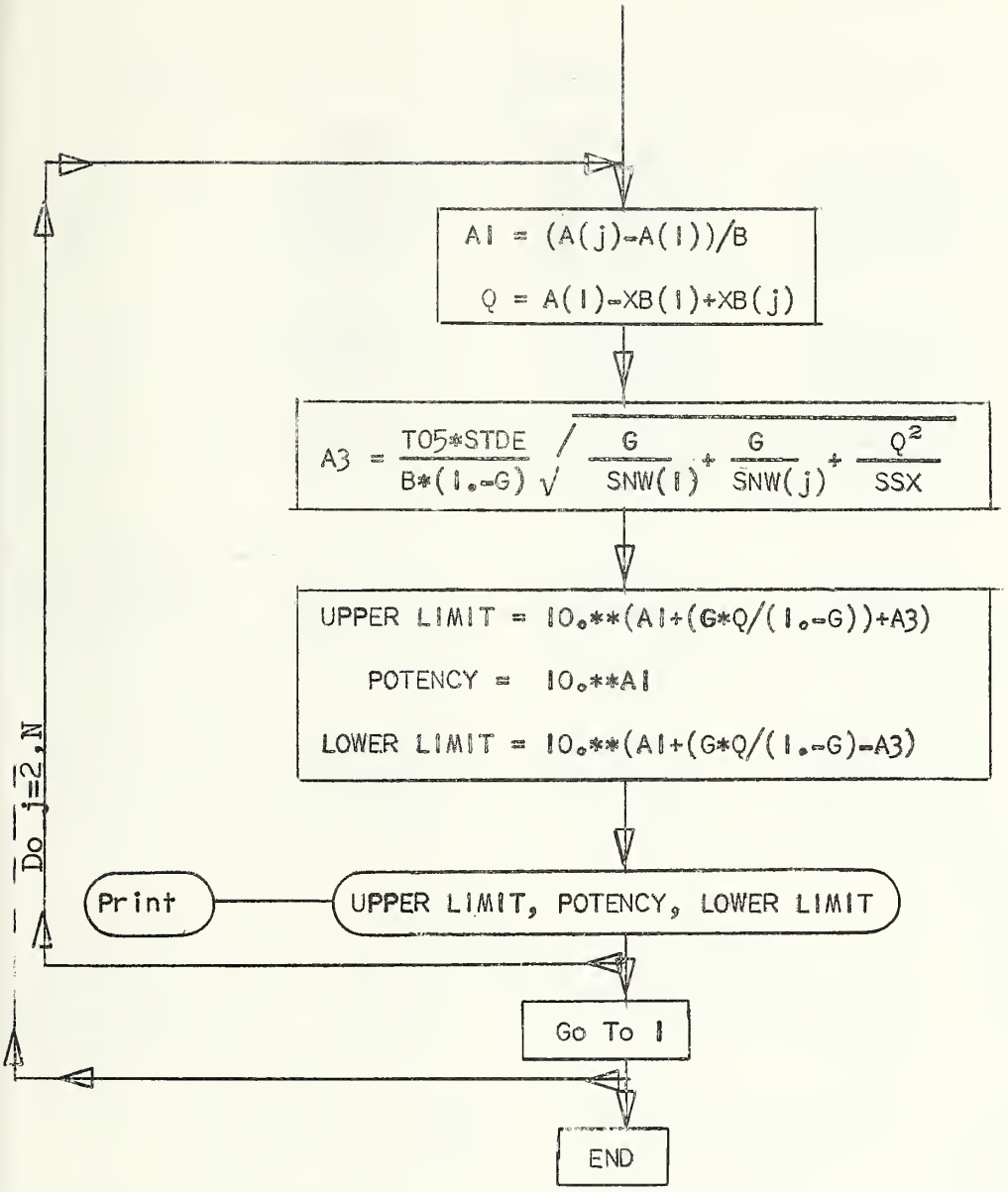



Do 11 j=1,N
Do 11 i=1,M
E = √LOG(1./Q²)
E = E-((a0+a1E+a2E²)/(1+b1E+b2E²+b3E³))
Z = .39894215*e^-(E²/2)
WN(i,j) = N(i,j)*Z²/(Q+(1.-Q)+(C(j)/(1.-C(j))))
P(i,j)-.5
Y(i,j) = 5.-E
Y(i,j) = 5.-E
Y(i,j) = 5.+E
CONTINUE
B = 0.
IT
Go To 1
SSY = 0.
SSX = 0.
SSXY = 0.
SSREG = 0.
SSNW = 0.
B2 = B
M = K(j)
XB(j) = 0.
YB(j) = 0.
SX(j) = 0.
SY(j) = 0.
SXY(j) = 0.
SNW(j) = 0.
XB(j) = XB(j)+X(i,j)*WN(i,j)
YB(j) = YB(j)+Y(i,j)*WN(i,j)
SNW(j) = SNW(j)+WN(i,j)
XB(j) = XB(j)/SNW(j)
YB(j) = YB(j)/SNW(j)


Do i=1,M
SX(j) = SX(j)+(X(i,j)-XB(j))**2*WN(i,j)
SY(j) = SY(j)+(Y(i,j)-YB(j))**2*WN(i,j)
SXY(j) = SXY(j)+(X(i,j)-XB(j))*(Y(i,j)-YB(j))*WN(i,j)
Do j=1,N
SSNW = SSNW+SNW(j)
SSX = SSX+SX(j)
SSY = SSY+SY(j)
SSXY = SSXY+SXY(j)
SSREG = SSREG+(SXY(j)*SXY(j))/SX(j)
B = SSXY/SSX
SREG = B*SSXY
IT = IT+1
|B2-B| ± .00005
M = K(j)
A(j) = YB(j)-B*XB(j)
Q = A(j)+B*X(i,j)
E = (Q-5.)/1.4142136
Q-5.
E = -E
E = .5
E = E*(E*(E*(E*(E*(E*D6+D5)+D4)+D3)+D2)+D1)
E = (1.-(1./(1.+E)**16.))*.5
Q-5.
E = .5-E
E = .5+E
Do j=1,N
Do i=1,M


Do j=1,N
Do i=1,M
Z = .39894215*e^-(E^2/2)
Y(i,j) = Q+((P(i,j)-E)/Z)
WN(i,j) = N(i,j)*Z^2/((1.-E)*(E+C(j)/(1.-C(j))))
Go To 12
TDF = 0
SSPAR = SSREG - SREG
Do j=1,N
SREG = SXY(j)*SXY(j)/SX(j)
TDF = TDF+K(j)-1
Print
j,K(j),SX(j),SY(j),SXY(j),SREG
RDF = 1
PDF = N-1
DDF = TDF-N
C(4) = 6.2816
C(3) = 5.5244
C(2) = 5.0000
C(1) = 4.4756
SPAR = SSPAR/PDF
SSDEV = SSY-SSREG
SDEV = SSDEV/DDF
FCAL = SREG/SDEV
Do j=1,N
Print
j,K(j),XB(j),YB(j),SWN(j),A(j)
Print
ANALYSIS OF LINEAR REGRESSION
Print
SOURCE OF VARIATION, DF, SS ,MS , Parameters
Print
Total ,TDF, SSY , ,F05
Print
REGRESSION ,RDF, SSREG, SSREG,FCAL
Print
DEVIATIONS ,DDF, SSDEV, SDEV, CH
Print
PARALLELISM ,PDF,SSPAR , SPAR, CH2


SSDEV-CH
Z = 1.00
T05 = 1.96
Z = √SDEV
STDEB = √Z²/SSX
G = T05*T05*Z/SREG
Print
IT,T05,STDE,STDEB,B
G-.05
G = 0.
F05-FCAL
Print
NON SIG REGRESSION
Go To 1
CH2-SSPAR
Print
LINES NOT PARALLEL
Go To 1
YB(i) = (10.**((C(1)-A(j))/B))/FAL
Do i=1,4
Print
IDE,j,YB(1),YB(2),YB(3),YB(4)
Do j=1,N


AI = (A(j)-A(I))/B
Q = A(I)-XB(I)+XB(j)
A3 = T05*STDE / B*(1.-G) √ G/SNW(I) + G/SNW(j) + Q²/SSX
UPPER LIMIT = 10.**(AI+(G*Q/(1.-G))+A3)
POTENCY = 10.**AI
LOWER LIMIT = 10.**(AI+(G*Q/(1.-G))-A3)
Print
UPPER LIMIT, POTENCY, LOWER LIMIT
Do j=2,N
Go To 1
END

IDENT	SET	DOSE	NET RESPONSE	CHECK
101	1	4.000000	16.666666	.000000
101	1	8.000000	40.000000	.000000
101	1	24.000000	65.000000	.000000
101	1	48.000000	78.500000	.000000
101	2	4.000000	2.666666	.000000
101	2	8.000000	5.714285	.000000
101	2	24.000000	28.444444	.000000
101	2	48.000000	45.777777	.000000
101	3	24.000000	30.000000	.000000
101	3	48.000000	39.000000	.000000

SET	NO.DOSES	SSX	SSY	SSXY	SSREG
1	4	63.918185	154.699630	98.430363	151.577140
2	4	32.391504	105.296120	58.167962	104.456760
3	2	2.649517	1.705918	2.125996	1.705918

SET	NO.DOSES	XBAR	YBAR	SNW	INTERCEPT
1	4	1.189489	5.080578	446.358190	3.172714
2	4	1.403076	4.419619	332.977760	2.169175
3	2	1.541593	4.609921	117.564690	2.137305

ANALYSIS OF LINEAR REGRESSION

SOURCE VARIATION	DF	SS	MS		
TOTAL	7	261.701660		7.710	F05
REGRESSION	1	254.583740	254.583740	257.035	FCAL
DEV REGRESSION	4	3.961850	.990462	9.490	CHI SQ
PARALLELISM	2	3.156070	1.578035	5.990	CHI SQ

IT	G	T05	STDE	STDEB	SLOPE
4	.015089	1.960000	1.000000	.100524	1.603936

IDENT	SET	LD30	LD50	LD70	LD90
101	1	6.490901	13.780042	29.254731	86.753000
101	2	27.414292	58.199950	123.557240	366.400940
101	3	28.697683	60.924561	129.341500	383.553880

IDENT	SET	UPPER LIMIT	POTENCY	LOWER LIMIT
101	2	.299413	.236770	.187233
101	3	.306311	.226182	.167013

IDENT	SET	DOSE	NET RESPONSE	CHECK
106	1	4.000000	34.666666	.000000
106	1	8.000000	71.500000	.000000
106	1	24.000000	86.153846	.000000
106	1	48.000000	96.500000	.000000
106	2	4.000000	17.333333	.000000
106	2	8.000000	40.571428	.000000
106	2	24.000000	77.333333	.000000
106	2	48.000000	89.333333	.000000
106	3	24.000000	60.000000	.000000
106	3	48.000000	69.000000	.000000

SET NO.	DOSES	SSX	SSY	SSXY	SSREG
1	4	45.918937	165.976230	84.454855	155.330740
2	4	45.486714	182.911510	91.017800	182.124380
3	2	2.645250	1.613487	2.065932	1.613487

SET NO.	DOSES	XBAR	YBAR	SNW	INTERCEPT
1	4	1.032328	5.572027	345.789870	3.623315
2	4	1.213743	5.367373	347.655920	3.076203
3	2	1.518727	5.359741	117.510200	2.492857

ANALYSIS OF LINEAR REGRESSION

SOURCE VARIATION	DF	SS	MS		
TOTAL	7	350.501220		7.710	F05
REGRESSION	1	335.137100	335.137100	117.256	FCAL
DEV REGRESSION	4	11.432620	2.858155	9.490	CHI SQ
PARALLELISM	2	3.931500	1.965750	5.990	CHI SQ

IT	G	T05	STDE	STDEB	SLOPE
3	.065720	2.776000	1.690607	.174325	1.887686

IDENT	SET	LD30	LD50	LD70	LD90
106	1	2.828115	5.361638	10.164779	25.599054
106	2	5.512247	10.450310	19.812055	49.894826
106	3	11.229392	21.289066	40.360545	101.644290

IDENT	SET	UPPER LIMIT	POTENCY	LOWER LIMIT
106	2	.794596	.513060	.319843
106	3	.467314	.251849	.130873

Code	Definition of Code Used in Printed Output
IDENT	Identification of the experiment or set of data.
SET	Identification of subset of data - subsets should be numbered consecutively for consistency of identifying numbers.
DOSE	The dose administered.
NET RESPONSE	The percent response adjusted for response in untreated control. (Note each subset may have a different control.)
CHECK	The proportion that responded in the untreated control. (Note each subset may have a different control.)
NO DOSES	The number of doses in each subset of data.
SSX	The weighted sum of squares of log dose corrected for its mean for each subset of data.
SSY	The weighted sum of squares of probit response corrected for its mean for each subset of data.
SSXY	The sum of the weighted cross products of log dose and probit response corrected for their means for each subset of data.
SSREG	The sum of squares of probit response attributable to linear regression of probits on log dose for each subset of data.
XBAR	The weighted mean of log dose ($\bar{x}$ = x bar), the mean may be negative if doses were less than one and "log factor" not used.
YBAR	The weighted mean of probit response.
SNW	The sum of the weighting coefficients for each subset of data.
INTERCEPT	The intercept of each regression line, or the value of Y (probit response) when x (log dose) is zero (0).
TOTAL	The sum of the SSY's (see above).
REGRESSION	The sum of SSXY squared and divided by the sum of the SSX's, which is the sum of squares attributable to a single slope for all lines (subsets of data).
DEV REGRESSION	The sum of the SSY's minus the sum of the SSREG's (see above, but do not confuse with REGRESSION), which is the weighted sum of squares of the deviations from parallel regression lines corrected for their means (i.e., the within deviations sum of squares).
PARALLELISM	The weighted sum of squares attributable to deviations from parallelism, which is the difference between the sum of SSREG's and the sum of SSXY's squared and divided by the sum of SSX's which is also the difference in sum of squares resulting from fitting one slope to all subsets of data and from fitting individual slopes for each subset of data. 1/
DF	Degrees of freedom associated with each source of variation.
F05	The critical value of F at 5% level with 1 and $\Sigma(k_i-1)$-N degrees of freedom, where K_i is the number of doses in each of the i subsets of data and N is the number of subsets of data or number of lines.
FCAL	The calculated F values from test of significance of the regression which is the ratio, REGRESSION/DEV REGRESSION mean squares.
CHI SQ	Chi square value at 5% probability level with $\Sigma(K_i-1)$-N degrees of freedom.
CHI SQ	Chi square value at 5% probability level with N-1 degrees of freedom.

1/ This definition is not entirely correct, but is probably the procedure that should be used. The reader interested in exactly how the "parallelism" sum of squares is calculated should examine the accompanying flow sheet or consult D. J. Finney's, Probit Analysis.

Code	Definition of Code Used in Printed Output
IT	The number of iterations required to obtain a difference between the last and the next to last estimates of the slope (B) of less than .00005. The number of iterations is counted starting from the second estimate of B as 1. The first estimate of B is obtained from a weighted linear regression of empirical probits on log dose, in which the net responses are used in the weighting coefficients rather than the proportions corresponding to the fitted provisional probits.
G	The precision of estimating the slope (will normally be the ratio of F05/FCAL). If G is less than .05, G is set equal to zero in calculating confidence limits.
T05	The critical value of students t at 5% probability level with $\Sigma(K_i-1)-N$ degrees of freedom which is used in setting confidence limits about the potencies. If data are heterogeneous T05 with $(K_i-1)-N$ degrees of freedom will appear here. If data are homogeneous 1.96 will appear here.
STDE	The standard error. If the data are homogeneous, the value 1.000000 will appear here. If the data are heterogeneous, the value $\sqrt{\text{DEV REGRESSION mean square}}$ will appear here.
STDEB	The standard error of the slope $= \sqrt{\text{STDE}^2/\ \text{SSX's}}$.
SLOPE	The common slope of all regression lines for the data composing "experiment."
LD-values	The LD-30, LD-50, LD-70, and LD-90 for each computed line. The computed lines may be obtained by plotting these values against 30, 50, 70, and 90% values on log-probit paper. The observed percent responses may also be plotted to give a graphical presentation of the numerical results.
UPPER LIMIT	The upper 95% confidence limit of potency.
POTENCY	The ratio of each succeeding LD-50 to the first LD-50 listed.
LOWER LIMIT	The lower 95% confidence limit of potency.

INSTRUCTIONS FOR PREPARING PARAMETER CARDS

One parameter card is required for each subset (up to five) of data. An additional parameter card for the entire set of data (experiment) is also required and precedes the first parameter card.

PARAMETER CARD FOR THE EXPERIMENT

The parameter card for the set (experimental) precedes all other cards and is referred to as the parameter card for the experiment.

Columns 1-4 Enter any 4-digit number to identify the experiment.

5-8 Enter any 4-digit number to identify the first subset of data (this field may be left blank but is useful in ordering cards).

9-11 Leave blank - not used.

12 Enter number, N, of subsets of data in the set (experiment). This number instructs the computer to read N subsets of data.

13-18 Enter Chi square value at 5% probability level with $\Sigma(K_i-1)-N$ degrees of freedom, where K_i is the number of doses in each of the i subsets of data, and N the number of subsets in the set (experiment). The Chi square value may be entered any place within this six digit field but decimal must be punched (see NOTE at end of instructions).

19-24 Enter Chi square value at 5% probability level with N-1 degrees of freedom. The Chi square value may be entered any place within this six digit field but decimal must be punched (see NOTE at end of instructions).

25-30 Enter critical value of Student's t with $\Sigma(K_i-1)-N$ degrees of freedom at 5% probability level. Decimal must be punched.

31-36 Enter critical value of F at 5% probability level with one and $\Sigma(K_i-1)-N$ degrees of freedom. Decimal point must be punched.

37-42 Enter LOG FACTOR - 10^i where i is an integer between 0 and 6 such that when the lowest dose is multiplied by 10^i it will be greater than unity. If all doses are already greater than unity the LOG FACTOR will be 10^0=000001. If the lowest dose is .00005 the LOG FACTOR will be 10^5 = 100000. This LOG FACTOR eliminates the possibility of a negative mean for log dose, which may confuse the biologist. Negative values will not effect the results nor interfere with the calculations.

PARAMETER CARD FOR EACH SUBSET

One parameter card precedes each subset of data.

Columns 1-4 Enter 4-digit number identifying the experiment (see above Column 1-4).

5-8 Enter 4-digit number identifying the subset of data. For consistency of output these subsets should be numbered consecutively starting with 1 (one).

9-10 Leave blank - not used.

11-12 Enter number, M, of doses or data cards (up to twenty) that comprise this subset of data. This number, M, instructs the computer to read the next M cards as data cards.

Columns 13-18 Enter number of animals observed in the untreated control (check) for this subset of data. Note that each subset may have a different value for the untreated control as well as a different number of doses.

19-24 Enter the number of animals that responded in the untreated control for this subset of data. Columns 13-24 may be left blank if no check was used or if no animals responded in the untreated control.

POTENCY PROBIT ANALYSIS

INSTRUCTIONS FOR PREPARING DATA CARDS

Enter data for each dose (dose used, number of animals treated, number of animals that responded) in the following card columns using one card for each dose.

Columns 1-4 Enter any 4-digit number to identify the experiment.

5-8 Enter any 4-digit number to identify the subset to which these data belong. For consistency in identification of "SET," the subsets should be numbered consecutively beginning with 1 (one).

9-14 Enter dose or concentration used. The dose may be entered any place within this field providing that decimal point is punched in its proper place. If no decimal point is punched, the computer will automatically place a decimal point between card columns 14 and 15.

15-20 Enter number of animals which received the dose listed. If no decimal point is punched, the program will automatically place a decimal point between card columns 20 and 21.

21-26 Enter the number of animals that responded to the dose listed on this card. If no decimal point is punched, the computer will automatically place a decimal point between card columns 26 and 27.

Assemble all data cards for each subset in either ascending or decending order of dose and precede each subset with its proper parameter card (see instructions for preparing parameter cards).

NOTE: If decimal points appear any place except between the designated fields, the decimal points should be punched. A punched decimal point, with FORMAT FORTRAN, will over-ride a FORMAT decimal point. Decimal points should not be punched for the numbers identifying the experiment and subsets of data, which are read as fixed rather than floating numbers (as I4 rather than F6.0).

POTENCY PROBIT ANALYSIS

Operating Instructions for IBM 1620

1. Clear Memory.

2. Load POTENCY PROBIT ANALYSIS program and FORMAT FORTRAN Subroutines.

3. Load data cards in following order:

    a. Parameter card for "experiment" or for all subsets of data.
    b. Parameter card for first subset of data.
    c. Data cards in ascending or decending order of dose for first subset.
    d. Parameter card for second subset of data.
    e. Data cards in ascending or decending order of dose for second subset.

    Additional subsets, up to five, may follow the first parameter card as outlined above.

4. Additional sets (experiments) may follow the first as indicated in (3) above.

5. Follow the last subset with 2 blank cards.

NOTES:

(1) Sense switches are not interrogated.
(2) When net response is negative, the program types out the doses for all subsequent subsets within this set, then goes to the next set of data ignoring the computations of the rest of the analysis. The decision to set the check for this subset in question to zero (0) and re-run the data is left to the experimenter or biometrician. It may be prudent for the biometrician to examine the reason(s) for different checks in each subset before submitting the data, and also to examine the numbers used in the checks for precision of estimating the proportion that responded in the untreated control. Abbott's formula assumes that the response in the untreated control is known without error.
(3) If a print-out "LINES NOT PARALLEL" or "NON SIG REGRESSION" occurs the remainder of the calculations are bypassed and the program procedes to the next set of data (experiment), ignoring the remaining calculations.
(4) Error F8 in the print-out may occur when the values are too large for the space allotted. When this occurs, the parameter and data cards should be examined for order and correctness of entry of the values.

POTENCY PROBIT ANALYSIS

Data Used in Example of Printed Output

Code	Set	Dose*	Number Observed	Number Responded
101	1	4	150	25
		8	200	80
		24	260	169
		48	200	157
101	2	4	75	2
		8	175	10
		24	225	64
		48	225	103
101	3	24	100	30
		48	100	39
Check		--	000	---
106	1	4	150	52
		8	200	143
		24	260	224
		48	200	193
106	2	4	75	13
		8	175	71
		24	225	174
		48	225	201
106	3	24	100	60
		48	100	69
Check		--	000	---

*Data supplied by courtesy of Dr. Harrie M. Taft from a time-mortality study. Boll weevils were exposed for 4, 8, 24, and 48 hours to treated foliage. Set one was from "immediately after application," sets two and three were from 24 and 48 hour weathering under artificial conditions. Purpose of study was to determine which of the 25 treatments (all recommended for boll weevil control) was the most toxic and which lost its effectiveness most rapidly. Loss in toxicity or loss in effectiveness is therefore 1.-potency, and its confidence limits may be similarly expressed. Treatment 101 lost 1.-.23677 = 76% of its effectiveness in 24 hours and 77% in 48 hours. Treatment 106 lost 1.-.51306 = 49% of its effectiveness in 24 hours and 75% in 48 hours. Treatment 106 was originally the most toxic (LD-50 = 5.36 hours) while treatment 101 the least toxic (LD-50 = 13.78 hours).